STRIFE, TRASH, BDSF, LRSG, or STMST modulator, such as a modulator identified by one of the exemplary screening assays described herein.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

EXAMPLES

EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN ITALY cDNA

In this example, the identification and characterization of the gene encoding human ITALY (also referred to as "TANGO 116") is described.

Isolation of the human ITALY cDNA

The invention is based, at least in part, on the discovery of a human gene encoding a novel protein, referred to herein as ITALY. In order to identify novel IL-10 family members, the Cytokine Cluster Program was used. This program is based on the observation that some genes were generated by gene duplication during evolution and the duplicated genes can be situated in close proximity to each other. Briefly, using a pair of human IL-10 primers for PCR, a positive BAC clone was identified from a BAC library (Research Genetics, Huntsville AL). The DNA from this BAC was extracted and made into a random-sheared genomic library. 4000 clones from this BAC were selected for sequencing. The resulting genomic sequences were then assembled into contigs and used to search proprietary and public data bases.

One genomic contig was found to match with four dbest (EST database) sequences (the forward and reverse sequences from two identical clones from human pregnant uterus: AA151652/AA151733 and AA151656/AA151736). Using one dbest clone (IMAGE ID# 503280) as a probe, a full-length cDNA clone was obtained by screening a PHA-stimulated PBL cDNA library (Clontech, Palo Alto, CA).

The sequence of the entire clone was determined and found to contain an open reading frame of 177 amino acids termed "Interleukin Ten Associated Locus Yang" or ITALY. Signal peptide algorithms predict that ITALY contains a signal peptide (amino acids 1-24 of SEQ ID NO:2). Cleavage of the putative signal peptide would result in the secretion of a 153 amino acid protein with a predicted molecular weight of 18 kilodaltons (kD).

The nucleotide sequence encoding the human ITALY protein is shown in Figure 1 and is set forth as SEQ ID NO:1. The full length protein encoded by this nucleic acid comprises about 177 amino acids and has the amino acid sequence shown in Figure 1

and set forth as SEQ ID NO:2. The coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3. Clone fthP116a, comprising the entire coding region of human ITALY was deposited with the American Type Culture Collection (ATCC®), Rockville, Maryland, on March 11, 1998, and assigned Accession No. 98960.

Analysis of Human ITALY

A BLAST search (Altschul et al. (1990) *J. Mol. Biol.* 215:403) of the nucleotide and protein sequences of human ITALY revealed that ITALY is similar to the following proteins: human MDA-7 (Accession No. Q13007; SEQ ID NO:9), mouse MDA-7 (SEQ ID NO:8; from a proprietary database), human IL-10 (Accession No.P22301; SEQ ID NO:5), mouse IL-10 (Accession No. P18893; SEQ ID NO:6), and viral IL-10 (Accession No.P03180; SEQ ID NO7). Each of these proteins is a member of the IL-10 family and contains both disulfide forming cysteine residues and an α-helical structure, although ITALY shares no greater than 32% identity with any of these cytokines. In particular, ITALY is 32% identical to human MDA-7 (Accession Number Q13007) (over nucleotides 157-561) and 22% identical to cat IL-10 (Accession Number P55029) (over nucleotides 214-543), at the amino acid level. An alignment of human ITALY and the above-described proteins is presented in Figure 2.

Tissue Distribution of ITALY mRNA

This Example describes the tissue distribution of ITALY mRNA, as determined by Northern blot hybridization and PCR.

Northern blot hybridizations with the various RNA samples were performed under standard conditions and washed under stringent conditions, i.e., 0.2xSSC at 65°C. A DNA probe corresponding to the coding region of ITALY (SEQ ID NO:3) was used. The DNA was radioactively labeled with ^{32}P-dCTP using the Prime-It kit (Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing human mRNA (MultiTissue Northern I and MultiTissue Northern II from Clontech, Palo Alto, CA) were probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

ITALY message was only detected by PCR from activated peripheral blood monocytes cDNA library (Clontech). No expression of ITALY was detected from any normal human tissue by northern analysis. The human tissues tested were: heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, lymph node, thymus, peripheral blood leukocyte, bone marrow, fetal liver, adrenal medulla, thyroid, adrenal cortex, testis, small intestine, stomach, prostate, uterus, and colon. However, a 2 Kb ITALY message was detected in brain tumors from unidentified cancer patients.

EXAMPLE 2: IDENTIFICATION AND CHARACTERIZATION OF HUMAN LOR-2 cDNA

In this example, the identification and characterization of the gene encoding human Lor-2 (*i.e.*, Lysyl Oxidase Related-2, also referred to as Myocardium Secreted Protein-18 or "MSP-18") is described.

Isolation of the human Lor-2 cDNA

The invention is based, at least in part, on the discovery of the human gene encoding Lor-2. Human Lor-2 was isolated from a cDNA library which was prepared from tissue obtained from subjects suffering from congestive heart failure. Briefly, a cardiac tissue sample was obtained from a biopsy of a 42 year old woman suffering from congestive heart failure. mRNA was isolated from the cardiac tissue and a cDNA library was prepared therefrom using art-known methods (described in, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989). Using a program which identifies the presence of signal peptides (Nielsen, H. *et al.* (1997) *Protein Eng.* 10:1-6) a positive clone was isolated.

The sequence of the positive clone was determined and found to contain an open reading frame. The nucleotide sequence encoding the human Lor-2 protein comprises about 2920 nucleic acids, and has the nucleotide sequence shown in Figures 3A-3B and set forth as SEQ ID NO:10. The protein encoded by this nucleic acid comprises about 753 amino acids, and has the amino acid sequence shown in Figures 4A-4C and set forth as SEQ ID NO:11.

Analysis of human Lor-2

A BLAST search (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide and protein sequences of human Lor-2 revealed that Lor-2 is similar to the following protein molecules: a human lysyl oxidase-related protein (Accession No. U89942; SEQ ID NO:16) having approximately 56.9% identity over amino acids 33-752 of Lor-2 (SEQ ID NO:11); and a second murine lysyl-oxidase related protein (Accession No.AF053368; SEQ ID NO:17) having approximately 92.6% identity over amino acids 1-753, *e.g.*, over the entire length) of Lor-2 (SEQ ID NO:11). (Identities were calculated using the LALIGN algorithm of Huang and Miller (1991) *Adv. Appl. Math.* 12:373-381).

The Lor-2 protein is predicted to have a signal peptide from amino acid residues 1-25 of SEQ ID NO:11. Accordingly, a mature Lor-2 protein is predicted to include amino acid residues 26-753 of SEQ ID NO:11. Lor-2 is also predicted to have 5 N-glycosylation sites, 8 protein kinase phosphorylation ("PKC") sites, 14 casein kinase II phosphorylation sites, 19 N-myristoylation sites, and 1 amidation site. Predicted N-

glycosylation sites are found, for example, from about amino acid 111-114, 266-269, 390-393, 481-484, and 625-628 of SEQ ID NO:11. Predicted PKC phosphorylation sites are found, for example, from about amino acid 97-99, 104-106, 221-223, 268-270, 352-354, 510-512, 564-566, and 649-651 of SEQ ID NO:11. Predicted casein kinase II phosphorylation sites are found, for example, from about amino acid 31-34, 68-71, 115-118, 120-123, 135-138, 330-333, 352-355, 377-380, 392-395, 411-414, 424-427, 493-496, 527-530, and 617-620 of SEQ ID NO:11. Predicted N-myristoylation sites are found, for example, from about amino acids 13-18, 116-121, 130-135, 273-278, 312-317, 359-364, 378-383, 403-408, 443-448, 451-456, 463-468, 470-475, 489-494, 506-511, 515-520, 521-526, 626-631, 661-666, and 746-751 of SEQ ID NO:11. A predicted amidation site is found, for example, from amino acid 117-180 of SEQ ID NO:11.

Moreover, Lor-2 has a 4 scavenger receptor cysteine-rich domains from amino acid residues 51-145, 183-282, 310-407, and 420-525 of SEQ ID NO:11. The third scavenger receptor cysteine-rich domain includes a speract receptor repeated domain signature from amino acid residues 312-349 of SEQ ID NO:11. Lor-2 further has a lysyl oxidase domain from residues 330-732 of SEQ ID NO:11. (See, for example, figures 7A-7B). Within the lysyl oxidase domain of Lor-2, there exists a fragment having significant homology to the lysyl oxidase putative copper-binding region, termed the "copper-binding talon". A ProSite consensus pattern describing the copper-binding talon is as follows: W-E-W-H-S-C-H-Q-H-Y-H (SEQ ID NO:18) (see also PROSITE documentation PDOC00716 and Krebs and Krawetz (1993) *Biochem. Biophys. Acta* 1202:7-12). Amino acid residues 601-701 of human Lor-2 (SEQ ID NO:11) have ~73% identity with this consensus sequence (8/11 residues) including each of the four conserved histidines, three of which are believed to be copper ligands residing within an octahedral coordination complex of lysyl oxidase.

Analysis of primary and secondary protein structures, as shown in Figure 6, was performed as follows: alpha, beta turn and coil regions, Garnier-Robson algorithm (Garnier *et al.* (1978) *J. Mol. Biol.* 120:97); alpha, beta, and turn regions, Chou-Fasman algorithm (Chou and Fasman (1978) *Adv. Enzymol. Mol.* 47:45-148); hydrophilicity and hydrophobicity plots, Kyte-Doolittle algorithm (Kyte and Doolittle (1982) *J. Mol. Biol.* 157:105-132); alpha amphipathic and beta amphipathic regions, Eisenberg algorithm (Eisenberg *et al.* (1982) *Nature* 299:371-374); flexible regions, Karplus-Schulz algorithm (Karplus and Schulz (1985) *Naturwissens-Chafen* 72:212-213); antigenic index, Jameson-Wolf algorithm (Jameson and Wolf (1988) *CABIOS* 4:121-136); surface probability plot, Emini algorithm (Emini *et al.* (1985) *J Virol* 55:836-839).

Prediction of the chromosomal location of Lor-2 – Electronic Mapping

To predict the chromosomal location of Lor-2, the Lor-2 nucleotide sequence of SEQ ID NO:10 was used to query, using the BLASTN program (Altschul S.F. et al. (1990) *J. Mol. Biol.* 215: 403-410) with a word length of 12 and using the BLOSUM62 scoring matrix, a database of human nucleotide sequences originating from nucleotide molecules (*e.g.*, EST sequences, STS sequences and the like) that have been mapped to the human genome. Nucleotide sequences which had been previously mapped to human chromosome 2 near the D2S145 marker (*e.g.*, having Accession Nos. AA191602 and R55706) were found to have high sequence identity to portions of the Lor-2 nucleotide sequence (3' UTR sequence) indicating that Lor-2 maps to the same chromosomal location. Moreover, it is predicted that allelic variants of Lor-2 will map the same chromosomal location and species orthologs of Lor-2 will map to loci syntenic with the human Lor-2 locus.

Confirmation and analysis of the chromosomal location of Lor-2 – PCR Mapping

The hLor-2 gene was mapped to human chromosome 2 (*i.e.*, 2p11-p13), which is syntenic to mouse chromosome 6, by PCR typing of the Genebridge (G4) radiation hybrid panel (Research Genetics, Inc., Huntsville, AL). Typing of the DNA and comparison to radiation hybrid map data at the Whitehead Institute Center for Genome Research (WICGR) tightly linked the hLor-2 gene to a region on human chromosome 2 between WI-5987 (13.9cR) and GCT1B4 (16.7cR).

The huLor-2 primers used in the PCR mapping studies were: forward - GCTTACCAAGAAACCCATGTCAGC (SEQ ID NO:20) and reverse - GGCAGTTAGTCAGGTGCTGC (SEQ ID NO:21). The radiation hybrid mapping studies were performed as follows: PCR reactions of radiation hybrid panels, GeneBridge 4 (Research Genetics, Inc., Huntsville, AL) were assembled in duplicate using an automated PCR assembly program on a TECAN Genesis. Each reaction consisted of: 5µl DNA template (10ng/µl), 1.5µl 10xPCR buffer, 1.2µl dNTPs (2.5mM), 1.1µl forward primer (6.6µM) 1.15µl reverse primer (6.6µM), and 5µl 1:75 platinum Taq. The reactions were thermocycled on a Perkin-Elmer 9600 for 95°C 10 minutes (for the platinum Taq), [95°C 40 sec, 52°C 40 sec, 72°C, 50 sec] 35X, 72°C, 5 minutes, 4°C hold. Resulting PCR products were run out on a 2% agarose gel and visualized on a UV light box.

The positive hybrids for the Genebridge 4 panel were submitted to the Whitehead Genome Center for placement in relation to a framework map.

Human Lor-2 mapped in close proximity to known genes including actin, gamma 2, smooth muscle, enteric ("ACTG2"), nucleolysin TIA1, semaphorin W ("SEMAW"), dysferlin ("DYSF"), docking protein 1 ("DOK1), glutamine-fructose-6-phosphate

transaminase 1 ("GFPT"), the KIAA0331 gene, deoxyguanosine kinase ("DGUOK"), the TSC501 gene, eukaryotic translation initiation factor 3, subunit 10 ("EIF3S1"), tachykinin receptor 1 ("TACR1"), tissue-type plasminogen activator ("PLAT") and dual specificity phosphatase 11 ("DUSP11"). Nearby disease mutations and/or loci include Alstrom syndrome ("ALMS1"), an autosomal recessively inherited syndrome characterized by retinal degeneration, obesity, diabetes mellitus, neurogenous deafness, hepatic dysfunction, and in some cases, late onset cardiomyopathy (see *e.g.*, Alstrom *et al.* (1959) *Acta Psychiat. Neurol. Scand.* 34 (suppl. 129):1-35; Alter and Moshang (1993) *Am. J. Dis. Child.* 147:97-99; Awazu *et al.* (1997) *Am. J. Med. Genet.* 69:13-16; Aynaci *et al.* (1995) *(Letter) Clin. Genet.* 48:164-166; Charles *et al.* (1990) *J. Med. Genet.* 27:590-592; Cohen and Kisch (1994) *Israel J. Med. Sci.* 30:234-236; Collin *et al.* (1997) *Hum. Molec. Genet.* 6:213-219; Collin *et al.* (1999) *(Letter) Clin. Genet.* 55:61-62; Connolly *et al.* (1991) *Am. J. Med. Genet.* 40:421-424; Goldstein and Fialkow (1973) *Medicine* 52:53-71; Macari *et al.* (1998) *Hum. Genet.* 103:658-661; Marshall *et al.* (1997) *Am. J. Med. Genet.* 73:150-161; Michaud *et al.* (1996) *J. Pediat.* 128:225-229; Millay *et al.* (1986) *Am. J. Ophthal.* 102:482-490; Rudiger *et al.* (1985) *Hum. Genet.* 69:76-78; Russell-Eggitt *et al.* (1998) *Ophthalmology* 105:1274-1280; Tremblay *et al.* (1993) *Am. J. Ophthal.* 115:657-665; Warren *et al.* (1987) *Am. Heart J.* 114:1522-1524 and Weinstein *et al.* (1969) *New Eng. J. Med.* 281:969-977), orofacial cleft 2 ("OFC2") (see *e.g.*, Carinci *et al.* (1995) *(Letter) Am. J. Hum. Genet.* 56:337-339; Pezzetti *et al.* (1998) *Genomics* 50:299-305 and Scapoli *et al.* (1997) *Genomics* 43:216-220) and Parkinson's disease 3 (see *e.g.*, Di Rocco *et al.* (1996) *Adv. Neurol.* 69:3-11 and Gasser *et al.* (1998) *Nature Genet.* 18:262-265). Additional information regarding Alstrom syndrome, orofacial cleft 2 and Parkinson disease 3 can be found collected under Accession Nos. 203800, 602966 and 602404, respectively, in the Online Mendelian Inheritance in Man ("OMIM™") database, the contents of which are incorporated herein by reference.

Moreover, the syntenic location on mouse chromosome 6 is near ovarian teratoma susceptibility 1 ("Ots-1"), disruption of corticosterone in adrenal cortex cells ("Cor"), brain protein 1 ("Brp1"), lymphocyte antigen 36 ("Ly36"), major liver protein 1 ("Lvp1"), cerebellar deficient folia ("cdf"), motor neuron degeneration 2 ("mnd2"), truncate ("tc") and faded ("fe"). Of particular interest are the Lor-2 neighbors Ots-1 and Cor, both of which a postulated to play a role in tumor susceptibility. The Ots-1 locus was identified by linkage analysis of female LT/Sv mice, a strain characterized by its abnormally high incidence of spontaneous ovarian teratomas, which are extremely rare for other mouse strains. Ots-1 was identified as the single major locus that increases the frequency of teratomas in a semidominant manner (Lee *et al.* (1997) *Cancer Res.* 57:590-593. Likewise, the cor locus was identified as being associated with a phenotype

of the AJ mouse strain (a strain susceptible to many neoplasms and infectious agents, presumably due to a deficiency in the prophylactic activities of endogenous glucocorticoids (*e.g.*, adrenalcortical corticosterone ("CS")) (Thaete *et al.* (1990) *Proc. Soc. Exp. Biol. Med.* 194:97-102). Accordingly, at least two loci in the near vicinity of mouse Lor-2 on chromosome 6 are associated with tumor susceptibility. Additional information regarding the Ots-1 and Cor loci can be found collected under Accession Nos. MGI:85864 and MGI:58993, respectively, in the Mouse Genomics Informatics database (available online), the contents of which are incorporated herein by reference. Likewise, information regarding the cdf locus, the mnd2 locus and the mouse Lor-2 gene (*i.e.*, the mouse ortholog of human Lor-2) can be found collected under Accession Nos. MGI:86274, MGI:97039 and MGI:1337004, respectively.

Additional markers (*e.g.*, EST markers, STS markers and the like) are set forth in Figure 8, as are the relative distances between markers.

Tissue Distribution of Lor-2 mRNA

Standard molecular biology methods (Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) were used to construct cDNA libraries in plasmid vectors from multiple human tissues. Individual cDNA clones from each library were isolated and sequenced and their nucleotide sequences were input into a database. The Lor-2 nucleotide sequence of SEQ ID NO:10 was used to query the tissue-specific library cDNA clone nucleotide sequence database using the BLASTN program (Altschul S.F. et al, (1990) J. Mol. Biol. 215: 403-410) with a word length of 12 and using the BLOSUM62 scoring matrix. Nucleotide sequences identical to portions of the Lor-2 nucleotide sequence of SEQ ID NO:10 were found in cDNA libraries originating from human endothelial cells, lymph node, bone, heart, neuron, and testes. Lor-2 nucleic acid sequences, fragments thereof, proteins encoded by these sequences, and fragments thereof as well as modulators of Lor-2 gene or protein activity may be useful for diagnosing or treating diseases that involve the tissues in which the Lor-2 mRNA is expressed. Likewise, when a similar analysis was performed using the Lor-2 sequence of SEQ ID NO:10 to query publicly available nucleotide sequence databases (*e.g.*, DBEST databases) using BLAST, sequences having high homology to the 3' untranslated region of human Lor-2 were identified in a Soares placenta normalized library and in Soares testis, B-cell and lung normalized libraries.

Northern blot hybridization with RNA samples was next performed under standard conditions and washed under stringent conditions, *i.e.*, 0.2 X SSC at 65°C. A DNA probe was radioactively labeled with ^{32}P-dCTP using the Prime-It kit (Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing various

tissue and cell line mRNAs were probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

On a human mRNA blot containing mRNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas, Lor-2 transcript (~3.0kb) was detected in all tissues tested but was most strongly detected in heart and placenta. Moreover, Lor-2 mRNA was strongly expressed in the G361 melanoma cell line and in the SW480 adenocarcinoma colon cell lines (as compared to expression in the HL60, HeLa53, K562, Molty, Raji, and SW480 cell lines (SW480 cell line expressing a 2.4kb transcript). Transcripts of 5kb and 2kb were also detected evidencing possible splice variants of Lor-2.

Testing of a larger panel of human tissues revealed the following expression levels. Expression levels were normalized to beta 2 expression.

TABLE I: hu Lor-2 Expression in Normal Tissues

Tissue Source	huLor-2 Expression	Beta 2 Expression	Relative Expression*
Lymph Node (MPI 79)	30.550	18.170	10.78
Lymph Node (NDR 173)	29.930	19.190	33.59
Heart (PIT 272)	26.145	18.170	57.06
Heart (PIT 273)	29.375	19.110	46.85
Lung (MPI 131)	29.650	19.480	50.04
Lung (NDR 185)	27.165	17.050	51.96
Kidney (MPI 58)	30.695	20.790	60.13
Spleen (MPI 360)	27.005	17.150	62.25
SK Muscle (MPI 38)	29.480	20.400	106.15
Fetal Liver (MPI 425)	30.065	20.520	75.85
Fetal Liver (MPI 133)	31.570	23.550	221.32
Tonsil (MPI 37)	29.480	17.890	18.64
Colon (MPI 383)	30.045	19.830	48.50
Brain (MPI 422)	30.525	22.220	181.65
Liver (MPI 75)	32.935	20.940	14.07
Liver (MPI 365)	31.060	18.770	11.35
Liver (MPI 339)	33.985	20.740	5.92
Liver (MPI 154)	32.000	19.970	13.74
Liver (NDR 206)	33.750	20.370	5.41

Liver (PIT 260)	32.705	18.970	4.23
CD14	26.945	17.190	66.49
Granulocytes	30.825	19.240	18.77
NHLH (resting)	36.595	19.920	1.10
NHLH (activated)	35.570	19.760	1.00
Liver Fibrosis (MPI 447)	29.320	18.300	27.67
Liver Fibrosis (NDR 190)	36.495	24.180	22.55
Liver Fibrosis (NDR 191)	30.105	19.770	44.63
Liver Fibrosis (NDR 192)	33.415	22.410	27.95
Liver Fibrosis (NDR 193)	30.795	19.830	28.74
Liver Fibrosis (NDR 204)	33.360	21.580	16.34
Liver Fibrosis (NDR 126)	31.900	21.180	34.18
Liver Fibrosis (NDR 113)	29.175	18.510	36.51
Liver Fibrosis (NDR 79)	30.870	20.390	40.22
Liver Fibrosis (NDR 112)	31.955	21.770	49.52
Liver Fibrosis (NDR 225)	30.645	20.350	45.89
Liver Fibrosis (NDR 141)	33.045	22.250	32.45

* NHLH activated used as reference sample

Next, Lor-2 expression levels were measured in a variety of tissue and cell samples using the Taqman™ procedure . The Taqman™ procedure is a quantitative, real-time PCR-based approach to detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA is generated from the samples of interest and serves as the starting materials for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) is included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer

sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product.

TABLE II: hu Lor-2 3' UTR Expression in Normal Human Tissues

Tissue Source	Relative Expression*
Prostate	2.5
Prostate	10.9
Liver	2.4
Liver	2.5
Breast	26.7
Breast	59.3
Skeletal Muscle	13.4
Skeletal Muscle	5.5
Brain	12.6
Brain	12.7
Colon	7.2
Colon	3.4
Heart	1.8
Heart	1.8
Ovary	1.8
Ovary	1.4
Kidney	1.0
Kidney	2.3
Lung	1.8
Lung	4.2
Vein	57.5
Vein	16.1

Tissue Source	Relative Expression*
Aorta	11.8
Testis	16.4
Testis	21.7
Thyroid	4.4
Thyroid	7.2
Placenta	73.3
Placenta	61.8
Fetal Kidney	87.7
Fetal Liver	10.0
Fetal Liver	64.7
Fetal Heart	14.4
Fetal Heart	70.8
Osteoblasts (undif.)	207.9
Osteoblasts (dif.)	128.0
Small Intestine	7.9
Cervix	86.5
Spleen	6.3
Esophagus	2.4
Thymus	1.4
Tonsil	1.7
Lymph node	3.1

* Kidney used as reference sample

The highest expression was observed in osteoblasts, cervix, kidney and placenta on the normal human tissue panel tested.

EXAMPLE 3: EXPRESSION OF LOR-2 mRNA IN CLINICAL TUMOR SAMPLES AND IN XENOGRAFT CELL LINES

In this example, RT-PCR was used to detect the presence of Lor-2 mRNA in various tumor and metastatic tissue samples as compared to normal tissue samples. RT-PCR was also used to detect the presence of Lor-2 mRNA in various xenograft cell lines. In breast tissue, Lor-2 mRNA was detected in 0/1 normal tissue samples as compared to 3/4 tumor clinical samples after 30 cycles of PCR. In xenograft cell lines isolated from breast tissue, Lor-2 mRNA was detected in 1/1 normal and 3/3 xenograft cell lines (cell lines MCF7, ZR75 and T47D). In lung tissue, Lor-2 mRNA was detected in 0/2 normal tissue samples as compared to 2/8 tumor tissue samples. In xenograft cell lines isolated from lung tissue, Lor-2 mRNA was detected in 0/5 xenograft cell lines after 30 cycles of PCR. In a second experiment performed with lung tissue, Lor-2 mRNA was detected in 2/2 normal and 8/8 tumor tissue samples, as well as in 5/5 xenograft cell lines (cell lines A549, H69, H125, H322 and H460) after 35 cycles of PCR. In colon tissue, Lor-2 mRNA was detected in 2/2 normal, 5/5 tumor and 5/5 metastatic samples, as well as in 7/7 xenograft cell lines (cell lines HCT116, HCT15, HT29, SW620, SW480, DLD1 and KM12) after 35 cycles of PCR. In liver tissue, LOR-2 mRNA was detected in 2/2 normal samples after 35 cycles of PCR. These data reveal that there exists a correlation between tumors and Lor-2 expression, at least in breast and lung tissues.

To further investigate this finding, Lor-2 mRNA levels were measured by quantitative PCR using the Taqman™ procedure as described above. The procedure was carried out on cDNA generated from various carcinoma samples and compared to normal counterpart tissue samples. In 5/7 breast carcinomas, a 2-86 fold upregulation of Lor-2 was observed as compared to 2/4 normal breast tissue samples. Likewise, in 4/7 lung carcinomas, a 2-17 fold upregulation was observed as compared to 3/4 normal lung tissue samples. The relative levels of Lor-2 mRNA detected in various normal, tumor and metastases samples are set forth in Table III.

Table III: hu Lor-2 Expression –
Taqman Analysis of Oncology Panel

Tissue Source	Relative Expression
Breast N	46.85
Breast N	18.96
Breast N	1.00
Breast N	11.75
Breast T	86.52
Breast T	37.27
Breast T	25.72
Breast T	60.76
Breast T	19.84
Breast T	22.24
Breast T	16.26
Lung N	9.32
Lung N	3.34
Lung N	1.65
Lung N	3.84
Lung T	4.26
Lung T	7.39
Lung T	9.13
Lung T	12.08
Lung T	6.48
Lung T	17.27
Lung T	28.15

Tissue Source	Relative Expression
Colon N	48.50
Colon N	4.94
Colon N	10.09
Colon N	4.94
Colon T	10.78
Colon T	10.89
Colon T	17.39
Colon T	10.82
Colon T	9.09
Colon T	26.63
Liver Met	10.93
Liver Met	10.30
Liver Met	12.25
Liver Met	12.91
Liver N	4.30
Liver N	3.69
Liver N	3.48
Liver N	5.41

These data reveal a significant upregulation of Lor-2 mRNA in at least breast and lung carcinomas. Moreover, there was a significant upregulation of Lor-2 expression in metastatic as compared to normal liver samples. Given that the mRNA for Lor-2 is expressed in a variety of tumors, with significant upregulation in carcinoma samples in comparison to normal samples, it is believed that inhibition of Lor-2 activity may inhibit tumor progression by affecting the adhesive properties of the tumor cells to surrounding tissues.

EXAMPLE 4: IDENTIFICATION OF MURINE STRIFE1 AND STRIFE2 cDNA

In this example, the isolation and characterization of the cDNA encoding murine STRIFE1 and STRIFE2 is described. STRIFE is a mouse gene which encodes a protein belonging to the TNFR family. Two splice forms have been identified, one that is predicted to be membrane bound (STRIFE1) and one that is secreted (STRIFE2).

STRIFE was identified as a TNFR homologue by a computer-based search of the public EST databases. More specifically, the murine STRIFE1 and STRIFE2 cDNA were identified by searching against a copy of the GenBank nucleotide database using the BLASTN™ program (BLASTN 1.3MP: Altschul et al., *J. Mol. Bio.* 215:403, 1990). Numerous clones that consisted mostly of 3' reads and some that were 5' reads within the 3' untranslated region were found by this search. The sequences were analyzed against a non-redundant protein database with the BLASTX™ program, which translates a nucleic acid sequence in all six frames and compares it against available protein databases (BLASTX 1.3MP:Altschul et al., *supra*). This protein database is a combination of the Swiss-Prot, PIR, and NCBI GenPept protein databases. Two clones (Accession Numbers AA036247 and AA003356) were obtained from the IMAGE consortium, and fully sequenced. The additional sequencing of AA036247 (T 127a; STRIFE1) extended the original EST by 623 nucleotides (see SEQ ID NO:22) and the further sequencing of AA003356 (T127b; STRIFE2) extended the original EST by 254 nucleotides (see SEQ ID NO:26).

A BLASTN™ search of the EST database revealed the following ESTs having significant homology to clone Accession Number AA036247:

EST Database hits	Species	Base Pairs Covered	% Identity	Coding?
Accession # AA495217	zebrafish	602-711	82	yes

A BLASTN™ search of the EST and nucleotide database revealed the following ESTs and nucleotides having significant homology to clone Accession Number AA003356:

EST Database hits	Species	Base Pairs Covered	% Identity	Coding?
Accession # AA686080	rat	297-367	64	yes
Accession # AA209382	human	150-210	67	yes
Accession # AA409240	mouse	284-319	80	yes
Accession # N91779	mouse	519-489	83	yes

EXAMPLE 5: TISSUE EXPRESSION OF THE STRIFE1 AND STRIFE2 GENE

Human I and mouse multiple tissue northern (MTN) blots (Clontech, Palo Alto, CA) containing 2μg of poly A+ RNA per lane were probed with a 750bp Eco*RI*/Not*I* fragment of the mouse STRIFE1 cDNA. The filters were prehybridized in 10 ml of Express Hyb hybridization solution (Clontech, Palo Alto, CA) at 68°C for 1 hour, after which 100 ng of ^{32}P labeled probe was added. The probe was generated using the Stratagene Prime-It kit, Catalog Number 300392 (Clontech, Palo Alto, CA). Hybridization was allowed to proceed at 68°C for approximately 2 hours. The filters were washed in a 0.05% SDS/2X SSC solution for 15 minutes at room temperature and then twice with a 0.1% SDS/0.1X SSC solution for 20 minutes at 50°C and then exposed to autoradiography film overnight at -80°C with one screen. The mouse tissues tested included: heart, brain, spleen, lung, liver, skeletal muscle, kidney, and testis. The human tissues tested included: heart, placenta, lung, liver, skeletal muscle, kidney, and pancreas.

There was a strong hybridization to both mouse and human heart, brain, and lung indicating that the approximately 4.4 kb STRIFE1 and STRIFE2 gene transcript is expressed in these tissues.

EXAMPLE 6: CHARACTERIZATION OF THE MURINE STRIFE1 AND STRIFE2 PROTEINS

STRIFE1 is approximately 981 nucleotides in length and has an open reading frame of 645 nucleotides that is predicted to encode a protein of 214 amino acids. STRIFE2 is approximately 655 nucleotides long with an open reading frame of 453 nucleotides predicted to encode a protein of 150 amino acids. Both clones have been subcloned into a variety of expression vectors including those for retroviral delivery and for expression in bacterial, yeast and mammalian cells.

BlastX searching of the protein database confirms the homology of this clone to various members of the TNFR family. The extracellular domains of STRIFE1 and STRIFE2 are approximately 40% identical to OX40. Importantly, a number of cysteine residues within the extracellular domains of STRIFE1 and STRIFE2 match the cysteine-rich domain signature of the TNFR/NGFR family (Prosite Accession PDOC00561). The program SignalP (Nielsen et al, 1997) predicts a 30 amino acid signal peptide at the very N-terminus of both STRIFE1 and STRIFE2 (i.e., residues 1-29 of SEQ ID NOs:22 and 6). The predicted molecular weight for STRIFE1 is approximately 23.55 kDa with the signal peptide and 20.34 kDa without the signal peptide which is presumed to be cleaved in the mature protein. There are no obvious motifs in the small intracellular

domain of STRIFE1. STRIFE2 is predicted to be 16.72 kDa with the signal peptide and 13.51 kDa without the signal peptide.

A FASTA search (described in Pearson W.R. & Lipman D. J. (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448, score matrix: PAM120) using the STRIFE1 protein sequence as a query, indicates that STRIFE1 is 85.7% identical to the human OAF065 receptor (Accession number W70387; described in PCT application number WO 98/38304, published on September 3, 1998) over amino acid residues 1-203. The results from this search are shown in Figures 15A-15B.

A FASTA search (described in Pearson W.R. & Lipman D. J. (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448, score matrix: PAM120) using the STRIFE1 nucleotide sequence as a query, indicates that STRIFE1 is 70.6% identical to the nucleic acid molecule encoding the human OAF065 receptor (Accession number V33362; described in PCT application number WO 98/38304, published on September 3, 1998) over nucleotide residues 65-981. The results from this search are shown in Figures 16A-16I.

Structure of the STRIFE1 and STRIFE2 Family proteins

An alignment of the amino acid sequences of murine STRIFE1, STRIFE2, and murine OX40 (Accession Number P47741; SEQ ID NO:30) is shown in Figures 14A-14B. Amino acid residues which are conserved between murine STRIFE1 and STRIFE2 family members are highlighted. The percent identity was calculated using the alignment generated using MegAlign™ sequence alignment software. The initial pairwise alignment step was performed using a Wilbur-Lipmann algorithm with a K-tuple of 2, a GAP penalty of 5, a window of 4, and diagonals saved set to = 4. The multiple alignment step was performed using the Clustal algorithm with a PAM 250 residue weight Table, a GAP penalty of 10, and a GAP length penalty of 10.

EXAMPLE 7: ISOLATION AND CHARACTERIZATION OF HUMAN TRASH cDNAs

In this example, the isolation of the gene encoding human TRASH (also referred to as "TANGO 118") is described.

A human TRASH cDNA was identified by searching against a copy of the GenBank nucleotide database using the BLASTN™ program (BLASTN 1.3MP: Altschul et al., *J. Mol. Bio.* 215:403, 1990). Numerous clones that consisted mostly of 3' reads and some that were 5' reads within the 3' untranslated region were found by this search. The sequences were analyzed against a non-redundant protein database with the BLASTX™ program, which translates a nucleic acid sequence in all six frames and compares it against available protein databases (BLASTX 1.3MP:Altschul et al., *supra*). This protein database is a combination of the Swiss-Prot, PIR, and NCBI GenPept

protein databases. One clone was obtained from the IMAGE consortium, and fully sequenced. The additional sequencing of this clone extended the original EST by 865 nucleotides further 5'. The cDNA for this clone is approximately 1346 nucleotides in length and has an open reading frame of 753 nucleotides that is predicted to encode a protein of 250 amino acids.

The original first pass sequence of the clone showed homology to human TNF-α using the BLASTX™ program. The nucleotide sequence and predicted amino acid sequence are shown in Figure 17 (corresponding to SEQ ID NO:31 and SEQ ID NO:32, respectively). The human TRASH protein (corresponding to amino acids 1-250 of the predicted amino acid sequence, SEQ ID NO:32) shows 21.0% identity to the human TNF-α protein (SEQ ID NO:43) and 24.6% identity to the human Tweak protein (SEQ ID NO:44) (see Figure 18). The percent identity was calculated using the alignment generated using MegAlign™ sequence alignment software. The initial pairwise alignment step was performed using a Wilbur-Lipmann algorithm with a K-tuple of 1, a GAP penalty of 3, a window of 5, and diagonals saved set to = 4. The multiple alignment step was performed using the Clustal algorithm with a PAM 250 residue weight Table, a GAP penalty of 10, and a GAP length penalty of 10.

This human TRASH protein contains a TNF signature motif, a TNF-like N-terminal signal transmembrane anchor for a type II membrane protein (corresponding to amino acids 1-44 of the predicted amino acid sequence, SEQ ID NO:32), two cysteine residues which may be disulfide linked (corresponding to amino acids 196 and 211 of the predicted amino acid sequence, SEQ ID NO:32), and two putative N-lined glycosylation sites (corresponding to amino acids 124 and 237 of the predicted amino acid sequence, SEQ ID NO:32). This human TRASH protein also contains three potential initiating ATG codons that would result in polypeptides of 250aa (SEQ ID NO:32), 233aa (SEQ ID NO:39), or 205aa (SEQ ID NO:41) encoding methionines at positions 1, 18, and 46 of SEQ ID NO:32, respectively. The predicted molecular weight for the 250aa TRASH is approximately 27.4kDa.

A BLASTN™ search of the EST database revealed the following ESTs having significant homology to clone Accession # AA481449:

EST Database hits	Species	Base Pairs Covered	% Identity	Coding?
Accession # AA405973	human	730-1224	97	yes
Accession # AA293679	human	1318-884	100	yes
Accession # AA394070	human	1318-891	100	yes
Accession # AA443577	human	939-515	99	yes

EXAMPLE 8: DISTRIBUTION OF TRASH mRNA IN HUMAN TISSUES

Northern Blot Analysis

The expression of TRASH was analyzed using Northern blot hybridization. For analysis of human TRASH, the 1.3kb insert of AA4481449 was used as a probe. The probe DNA was radioactively labeled with ^{32}P-dCTP using the Prime-It kit™ (Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing mRNA (human MTNI and MTNII and murine embryo MTN from Clontech, Palo Alto, CA;) were probed in ExpressHyb™ hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

Expression was found predominantly in the peripheral blood leukocytes where a message of approximately 1.5 kb transcript was observed. In addition, a 1.5kb transcript was observed in the spleen and lymph nodes and lung. A slightly larger message, approximately 1.7kb was observed in the colon, spleen, and kidney.

EXAMPLE 9: IDENTIFICATION AND CHARACTERIZATION OF HUMAN AND MURINE BDSF cDNA

In this example, the identification and characterization of the genes encoding human BDSF-1 (also referred to as "TANGO 122" and "hT122") and murine BDSF-1 (also referred to as "mT122") are described.

Isolation of the human BDSF cDNA

The invention is based, at least in part, on the discovery of a human gene encoding a novel brain-derived signaling factor protein, referred to herein as BDSF. The methodology used to isolate the human BDSF-1 gene takes advantage of the fact that molecules such as BDSF-1 have an amino terminal signal sequence which directs certain secreted and membrane-bound proteins through the cellular secretory apparatus.

The human BDSF-1 mRNA was identified by screening of a human fetal brain cDNA library. This library was prepared using mRNA purchased from Clontech, Palo Alto (Cat. no, 6573-1). A signal trap cDNA library was prepared by ligating random primed double stranded cDNA into the expression vector, ptrAP1, resulting in fusions of cDNAs to the reporter, alkaline phosphatase (AP). DNAs from individual clones from this library were prepared by standard techniques and transfected into human embryonic kidney fibroblasts (293T cells). After 48 hours, cell supernatants were collected and assayed for AP activity. Clones giving rise to detectable AP activity in the supernatants of transfected cells were analyzed further by DNA sequencing and the novel clones subjected to further DNA sequencing.

The nucleotide sequence encoding the human BDSF-1 protein is shown in Figures 19A-19B and is set forth as SEQ ID NO:45. The full length protein encoded by this nucleic acid is comprised of about 244 amino acids and has the amino acid sequence shown in Figures 19A-19B and set forth as SEQ ID NO:46. The coding portion (open reading frame) of SEQ ID NO:45 is set forth as SEQ ID NO:47. Notable features of the human BDSF-1 protein include a signal peptide (about amino acids 1-25 of SEQ ID NO:46), an Ig-like domain (about amino acids 41-129 of SEQ ID NO:46) and two conserved cysteine residues (about amino acids 48 and 127 of SEQ ID NO:46). A clone, comprising the entire coding region of human BDSF-1 has been deposited with the American Type Culture Collection (ATCC), Rockville, Maryland, on May 15, 1998 as Accession No. 98756.

Isolation of the murine BDSF cDNA

The murine BDSF-1 gene was identified from a murine choroid plexus cDNA library. More specifically, a murine choroid plexus cDNA library was plated out and colonies picked into 96 well plates. The colonies were cultured, plasmids were prepared from each well, and several of the inserts were sequenced. The nucleotide sequences were compared against the human BDSF-1 nucleotide sequence. Upon review of the results from this sequence comparison, a murine BDSF-1 gene was obtained.

The nucleotide sequence encoding the murine BDSF-1 protein is shown in Figures 20A-20C and is set forth as SEQ ID NO:50. The full length protein encoded by this nucleic acid is comprised of about 251 amino acids and has the amino acid sequence shown in Figures 20A-20C and set forth as SEQ ID NO:51. The coding portion (open reading frame) of SEQ ID NO:50 is set forth as SEQ ID NO:52. Notable features of the murine BDSF-1 protein include a signal peptide (about amino acids 1-24 of SEQ ID NO:51), an Ig-like domain (about amino acids 40-128 of SEQ ID NO:51) and two conserved cysteine residues (about amino acids 47 and 126 of SEQ ID NO:51).

Analysis of Human and Murine BDSF

A BLAST search (Altschul et al. (1990) J. Mol. Biol. 215:403) of the nucleotide and protein sequences of human BDSF and murine BDSF has revealed that BDSF does not display significant homology to other proteins except for distant similarity to other Ig-like domain containing proteins. An alignment of human BDSF and murine BDSF, however, indicated that these proteins are 77.4% identical overall and 90.5% identical over the first 211 amino acids. The alignment of human BDSF-1 and murine BDSF-1 is presented in Figure 21. The alignment of human BDSF-1 and murine BDSF-1 was performed using the ALIGN program. When utilizing the ALIGN program for

comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 was used.

A search using the amino acid sequence of SEQ ID NO:46 was performed against the HMM database resulting in the identification of an Ig-like domain in the amino acid sequence of SEQ ID NO:46 and a score of 22.43 against the Ig family HMM Accession PF00047. The results of the search are set forth below:

Score: 22.43 SEQ ID NO:46: aa41-129 HMM: aa1-47 (SEQ ID NO:57)

```
        GqsVTLTCmVs.fhPpdYt.IwWY.rNgqpi.................
  41    GQNVEMSCAFQSGSASVYLEIQWWFLRGPEDLDPGAEGAGAQVELLPDR

        ....................tLtInsWqyEDsGtYwCmV
  90    DPDSDGTKISTVKVQGNDISHKLQISKVRKKDEGLYECRV
```

In another example, a search was performed using the amino acid sequence of SEQ ID NO:51 against the HMM database resulting in the identification of an Ig-like domain in the amino acid sequence of SEQ ID NO:51 and a score of 22.43 against the Ig family HMM Accession PF00047. The results of the search are set forth below.

Score: 22.43 SEQ ID NO:51 aa40-128 HMM: aa1-47 (SEQ ID NO:57)

```
        GqsVTLTCmVs.fhPpdYt.IwWY.rNgqpi.................
  40    GQNVEMSCAFQSGSASVYLEIQWWFLRGPEDLEQGTEAAGSQVELLPDR

        ....................tLtInsWqyEDsGtYwCmV
  89    DPDNDGTKISTVKVGGNDISHKLQISKVRKKDEGLYECRV
```

Accordingly, in one embodiment of the invention, a BDSF protein is a human BDSF-1 protein having an Ig-like domain at about amino acids 41-129 of SEQ ID NO:46. Such an Ig-like domain has the amino acid sequence of SEQ ID NO:55.

Expression of human BDSF

The expression of human BDSF was analyzed using Northern blot hybridization and a probe specific for human BDSF. The DNA was radioactively labeled with ^{32}P-dCTP using the Prime-It-kit (Stratagene, La Jolla, CA) according to the instructions of

the supplier. Filters containing human mRNA (Multi-Tissue Northern I, Multi-Tissue Northern II and Multi-Tissue Northern III from Clontech, Palo Alto, CA) were probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations. In addition, filters containing human brain mRNA (Brain-Subregion Blot from Clontech) were also probed for human BDSF-1 expression.

Results of Northern blot hybridization indicate that human BDSF is expressed as an approximately 5.0 kilobase transcript in all tissues (spleen, lymph node, thymus, peripheral blood leukocytes, bone marrow, stomach, thyroid, spinal cord, trachea, adrenal, testis, small intestine, heart, placenta, lung, liver, kidney and pancreas). BDSF mRNA expression was also observed in human fetal liver.

The highest level of human BDSF expression was found in the adult brain where the pattern of transcripts was also different than for the other tissues. In the brain, major transcripts of 2.6 kb, 3.2 kb and 6.5 kb were observed. This pattern of BDSF mRNA transcripts in the brain is found in the sub-regions of the brain including amygdala, caudate nucleus, hippocampus, substania nigra, sub-thalamate nucleus and thalamus, but not in corpus callosum.

EXAMPLE 10: IDENTIFICATION AND CHARACTERIZATION OF HUMAN LRSG-1 cDNA

In this example, the identification and characterization of the gene encoding human LRSG-1 (also referred to as human "TANGO 124") is described.

Isolation of the human LRSG-1 cDNA

The invention is based, at least in part, on the discovery of a human gene encoding a novel leucine-rich repeat containing protein, referred to herein as LRSG-1. Human astrocytes (obtained from Clonetics Corporation; San Diego, CA) were expanded in culture with Astrocyte Growth Media (AGM; Clonetics) according to the recommendations of the supplier. When the cells reached ~80-90% confluence, they were stimulated with 200 units/ml Interleukin 1-Beta (Boehringer Mannheim) and cyclohexamide (CHI; 40 micrograms/ml) for 4 hours. Total RNA was isolated using the RNeasy Midi Kit (Qiagen; Chatsworth, CA), and the poly A+ fraction was further purified using Oligotex beads (Qiagen).

Three micrograms of poly A+ RNA were used to synthesize a cDNA library using the Superscript cDNA Synthesis kit (Gibco BRL; Gaithersburg, MD). Complementary DNA was directionally cloned into the expression plasmid pMET7 using the SalI and NotI sites in the polylinker to construct a plasmid library. Transformants were picked and grown up for single-pass sequencing. Additionally, astrocyte cDNA was ligated into the SalI/NotI sites of the ZipLox vector (Gibco BRL)

for construction of a lambda phage cDNA library. A clone (jthxe016d10) that encoded a protein with limited homology to decorin, insulin-like growth factor binding protein and biglycan was identified. Full sequencing of the clone demonstrated that it contained an ~2.8kb insert with a single large open reading frame predicted to encode a 673 amino acid transmembrane protein.

The nucleotide sequence encoding the human LRSG-1 protein is shown in Figures 22A-22E and is set forth as SEQ ID NO:58. The full length protein encoded by this nucleic acid is comprised of about 673 amino acids and has the amino acid sequence shown in Figures 22A-22E and set forth as SEQ ID NO:59. The coding portion (open reading frame) of SEQ ID NO:58 is set forth as SEQ ID NO:60. Clone jthxe016d10, comprising the entire coding region of human LRSG-1 has been deposited with the American Type Culture Collection (ATCC), Manassas, Virginia on March 12, 1998 as accession Number 98695.

Notable features of the human LRSG-1 protein include a signal peptide (about amino acids 1-23 of SEQ ID NO:59), a transmembrane domain (about amino acids 576-599 of SEQ ID NO:59) an EGF-like domain (about amino acids 409-441) and a fibronectin type III-like domain (about amino acids 460-535 of SEQ ID NO:59). Th human LRSG-1 protein further includes a leucine-rich region (about amino acids 77-309 of SEQ ID NO:59) which includes at least 7 leucine-rich repeats (about amino acids 77-309, 101-123, 125-147, 149-171, 217-238, 240-263, and 289-309 of SEQ ID NO:59).

Analysis of Human LRSG-1

A BLAST search (Altschul *et al.* (1990) J. Mol. Biol. 215:403) of the nucleotide and protein sequences of human LRSG-1 has revealed that LRSG-1 has structural similarities with both platelet glycoprotein V precursor (GPV) (SwissProt Accession No. P40197; SEQ ID NO:61) and insulin-like growth factor binding protein complex acid labile chain precursor (ALS) (SwissProt Accession No. O02833; SEQ ID NO:62). Each of these proteins is a leucine-rich repeat containing protein although LRSG-1 shares no greater than 30% identity with any of these LRR-containing proteins. An alignment of human LRSG-1 and the above-described proteins is presented in Figures 23A-23B.

Human LSRG-1 was analyzed for the presence of an FN type III - like domain. An alignment with an FN type III - like consensus sequence is shown below for amino acid residues 460-535 of SEQ ID NO:59:

```
FN type III - like consensus  PsPPrNLrvtdITpTSItVSWtPPe..gNgpItgYr
                              P+    L +++++PTS++V  ++     +  +++ R
LRSG-1                        PPRSLTLGIEPVSPTSLRVGLQRYLQGSSVQLRSLR

FN type III - like consensus  IqYRWpvNdne..WnEfnVPrttnsYTItnLrPGTeYeFRV
                              ++YR +  +++   +++++P +  +YT+T LRP+ +Y++ V
LRSG-1                        LTYR-NLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCV
```

Expression of LRSG-1

The expression of LRSG-1 was analyzed using Northern blot hybridization. A 579 base pair (bp) DNA fragment from the N-terminal portion of the coding region was generated using PCR for use as a probe. The DNA was radioactively labeled with ^{32}P-dCTP using the Prime-It-kit (Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing human mRNA (Multi-Tissue Northern I and Multi-Tissue Northern II from Clontech, Palo Alto, CA) were probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

Results of Northern blot hybridization indicate that LRSG-1 is expressed as an approximately 3.0 kilobase transcript in all tissues (spleen, thymus, prostate, testes, ovary, small intestine, colon, heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas) with the exception of peripheral blood leukocytes. The highest levels of LRSG-1 were found in placenta, kidney and testis.

EXAMPLE 11: IDENTIFICATION AND CHARACTERIZATION OF MURINE LRSG-1 cDNA

In this example, the identification and characterization of the gene encoding murine LRSG-1 (also referred to as murine "TANGO 124") is described.

Using sequence information from the cloning of human LRSG-1, a murine homologue was identified.

The nucleotide sequence encoding the murine LRSG-1 protein is shown in Figure 24A-24E and is set forth as SEQ ID NO:67. The full length protein encoded by this nucleic acid is comprised of about 673 amino acids and has the amino acid sequence shown in Figure 24A-24E and set forth as SEQ ID NO:68. The coding portion (open reading frame) of SEQ ID NO:67 is set forth as SEQ ID NO:69.

Notable features of the murine LRSG-1 protein include a signal peptide (about amino acids 1-24 of SEQ ID NO:68), a transmembrane domain (about amino acids 577-600 of SEQ ID NO:68) an EGF-like domain (about amino acids 410-442) and a fibronectin type III-like domain (about amino acids 461-636 of SEQ ID NO:68). The murine LRSG-1 protein further includes a leucine-rich region (about amino acids 78-310 of SEQ ID NO:68) which includes at least 7 leucine-rich repeats (about amino acids 78-310, 102-124, 126-148, 150-172, 218-238, 241-264, and 290-310 of SEQ ID NO:68). An Fn type III-like domain was identified at about residues 461-536 of SEQ ID NO:68

Figures 25A-25C set forth a pairwise alignment of the amino acid sequences of human and murine LRSG-1. The alignment was generated using the ALIGN algorithm, version 2, which is part the GCG software package. ALIGN (Myers and Miller, CABIOS (1989)) calculates a global alignment of two sequences and is part the GCG software package. The alignment was generated using a PAM120 scoring matrix and gap penalties of -12/-4. Murine LRSG-1 is 83.2% identical to human LRSG-1 (global alignment score: 2910).

EXAMPLE 12: CHROMOSOMAL MAPPING OF LRSG-1 cDNA

PCR primers were designed based of the coding sequence of human LRSG-1 and used to generate probes for chromosomal mapping. LRSG-1 was found to map to chromosome 16 between markers WI-7742 and WI-3061 (46-52cM).

EXAMPLE 13: IDENTIFICATION AND CHARACTERIZATION OF STMST-1 cDNAs

In this example, the identification and characterization of the genes encoding human STMST-1 and STMST-2 (also referred to as "TANGO123a" and "TANGO 123c", respectively) is described.

Isolation of the human STMST cDNAs

In order to identify novel secreted and/or membrane-bound proteins, a program termed 'signal sequence trapping' was utilized to analyze the sequences of several cDNAs of a cDNA library derived from bronchial epithelial cells which had been stimulated with the cytokine, TNFα. This analysis identified a partial human clone having an insert of approximately 231 kb containing a protein-encoding sequence of approximately 225 nucleotides capable of encoding approximately 75 amino acids of STMST (*e.g.*, the start met through residue 74 of, for example, SEQ ID NO:71). This cDNA was used to re-screen the library. Two full-length cDNA clones were isolated. Sequencing of these clones revealed the nucleotide sequences of human STMST-1 and STMST-2.

The nucleotide sequence encoding the human STMST-1 protein is shown in Figures 27A-27B and is set forth as SEQ ID NO:70. The full length protein encoded by this nucleic acid is comprised of about 297 amino acids and has the amino acid sequence shown in Figures 27A-27B and set forth as SEQ ID NO:71. The coding portion (open reading frame) of SEQ ID NO:70 is set forth as SEQ ID NO:72.

The nucleotide sequence encoding the human STMST-2 protein is shown in Figures 28A-28C and is set forth as SEQ ID NO:73. The full length protein encoded by this nucleic acid is comprised of about 609 amino acids and has the amino acid sequence shown in Figures 28A-28C and set forth as SEQ ID NO:74. The coding portion (open reading frame) of SEQ ID NO:73 is set forth as SEQ ID NO:75.

Analysis of Human STMST-1 and STMST-2

Examination of the cDNA sequences depicted in Figures 27A-27B and 28A-28C showed that they were likely encoded by alternatively spliced mRNAs derived from the same gene. Thus, the amino acid sequence of STMST-1 diverges from that of STMST-2 at about amino acid residue 263 of SEQ ID NO:71 or SEQ ID NO:74. The amino acid sequence of STMST-1 lacks the extensive cytoplasmic domain of STMST-2.

A BLAST search (Altschul *et al.* (1990) J. Mol. Biol. 215:403) of the nucleotide sequence of human STMST-2 has revealed that STMST-2 is significantly similar to a protein identified as protein A-2 (human A-2, Accession No. U47928; murine A-2, Accession No. AC002393) which were sequenced as part of the sequencing of human chromosome 12p13 and mouse chromosome 6, respectively. The human A-2 protein appears to be one of a family of alternatively-spliced gene products which further includes protein A-1 (Accession No. U47925) as well as A-3 (Accession No. U47929). The A-2 proteins, like the STMST proteins of the present invention, include many features indicative of the G protein-coupled receptor family of proteins.

For instance, the STMSTs of the present invention contain conserved cysteines found in the first 2 extracellular loops (prior to the third and fifth transmembrane domains) of most GPCRs (cys 83 and cys 161 of SEQ ID NO:71 or SEQ ID NO:74). A highly conserved asparagine residue in the first transmembrane domain is present (asn25 in SEQ ID NO:71 or SEQ ID NO:74). Transmembrane domain two of the STMST proteins contains a highly conserved leucine (leu49 of SEQ ID NO:71 or SEQ ID NO:74). The two cysteine residues are believed to form a disulfide bond that stabilizes the functional protein structure. A highly conserved tryptophan and proline in the fourth transmembrane domain of the STMST proteins is present (trp135 and pro 145 of SEQ ID NO:71 or SEQ ID NO:74). The third cytoplasmic loop contains 49 amino acid residues and is thus the longest cytoplasmic loop of the three, characteristic of G protein coupled receptors. Moreover, a highly conserved proline in the sixth transmembrane domain is present (pro260 of SEQ ID NO:71 and SEQ ID NO:74). The proline residues in the fourth, fifth, sixth, and seventh transmembrane domains are thought to introduce kinks in the alpha-helices and may be important in the formation of the ligand binding pocket. Furthermore, the conserved (in the second cytoplasmic loop) HRM motif found in almost all Rhodopsin family GPCRs is present in the STMST proteins of the instant invention (his107, arg108, met109 of SEQ ID NO:71 or SEQ ID NO:74). (The arginine of the HRM sequence is thought to be the most important amino acid in GPCRs and is invariant). Moreover, an almost invariant proline is present in the seventh transmembrane domain of STMST-2 (pro294 of SEQ ID NO:74).

As such, the STMST family of proteins, like the A-2 family of proteins, are referred to herein as G protein-coupled receptor-like proteins.

STMST-1 is also predicted to contain the following sites: cAMP and cGMP-dependent protein kinase phosphorylation site at amino acid residues 225-228 (KRRS; SEQ ID NO:90); Protein kinase C phosphorylation sites at residues 153-155 (SER) and at residues 290-292 (SSR); Casein kinase II phosphorylation sites at residues 228-231 (SSID; SEQ ID NO:91) and at residues 291-294 (SRQD; SEQ ID NO:92); N-myristoylation sites at residues 9-14 (GSAVGW; SEQ ID NO:93), residues 169-174 (GLGFGV; SEQ ID NO:94), residues 181-186 (GGSVAM; SEQ ID NO:95), residues 187-192 (GVICTA; SEQ ID NO:96), residues 232-237 (GSEPAK; SEQ ID NO:97), and at residues 244-249 (GLVTTI; SEQ ID NO:98); Amidation site at residues 223-226 (QGKR; SEQ ID NO:99).

Likewise, STMST is predicted to contain the following sites: cAMP- and cGMP-dependent protein kinase phosphorylation sites at amino acid residues 225-228 (KRRS; SEQ ID NO:100), residues 393-396 (RRFS; SEQ ID NO:101), residues 436-439 (RRAS; SEQ ID NO:102), and at residues 453-456 (RRRS; SEQ ID NO:103); Protein kinase C phosphorylation sites at residues 253-255 (SER), residues 268-270 (SLR),

residues 392-394 (TRR), residues 462-464 (SLR), residues 482-484 (SPR), and at residues 560-562 (SLR); Casein kinase II phosphorylation sites at residues 228-231 (SSID; SEQ ID NO:104), residues 324-327 (SDDE; SEQ ID NO:105), residues 328-331 (TSLE; SEQ ID NO:106), residues 364-367 (SALE; SEQ ID NO:107), residues 396-399 (SHDD; SEQ ID NO:108), residues 417-420 (SGED; SEQ ID NO:109), residues 466-469 (SALD; SEQ ID NO:110), residues 506-509 (TAFE; SEQ ID NO:111), residues 568-571 (SWGE; SEQ ID NO:112), and at residues 590-593 (SPSE; SEQ ID NO:113); Tyrosine kinase phosphorylation site at residues 342-348 (RSLDYGY; SEQ ID NO:114); N-myristoylation sites at residues 9-14 (GSAVGW; SEQ ID NO:115), residues 169-174 (GLGFGV; SEQ ID NO:116), residues 181-186 (GGSVAM; SEQ ID NO:117), residues 187-192 (GVICTA; SEQ ID NO:118), residues 232-237 (GSEPAK; SEQ ID NO:119), residues 244-249 (GLVTTI; SEQ ID NO:120), residues 531-536 (GADPGE; SEQ ID NO:121), residues 564-569 (GLSASW; SEQ ID NO:122), residues 573-578 (GGLRAA; SEQ ID NO:123), and at residues 579-584 (GGGGST; SEQ ID NO:124); Amidation site at residues 223-226 (QGKR; SEQ ID NO:125).

The following table depicts an alignment of the transmembrane domain of 5 GPCRs. The conserved residues described herein are indicated by asterisks. An alignment of the transmembrane domains of 44 representative GPCRs can be found online.

ALIGNMENT OF:

thrombin	(6.)	human	P25116	(SEQ ID NO:80)
rhodopsin	(19.)	human	P08100	(SEQ ID NO:81)
m1ACh	(21.)	rat	P08482	(SEQ ID NO:82)
IL-8RA	(30.)	human	P25024	(SEQ ID NO:83)
octopamine	(40.)	Drosophila melanogaster	P22270	(SEQ ID NO:84)

```
TM1

                              *
 6.    102 TLFVPSVYTGVFVVSLPLNIMAIVVFILKMK        132  (SEQ ID NO:137)
19.     37 FSMLAAYMFLLIVLGFPINFLTLYVTVQHKK         67  (SEQ ID NO:138)
21.     25 VAFIGITTGLLSLATVTGNLLVLISFKVNTE         55  (SEQ ID NO:139)
30.     39 KYVVIIAYALVFLLSLLGNSLVMLVILYSRV         69  (SEQ ID NO:140)
40.    109 ALLTALVLSVIIVLTIIGNILVILSVFTYKP        139  (SEQ ID NO:141)
                              |
           11111111111111111111111111111111
           33333333444444444455555555566666
           23456789012345678901234567890123

TM2

                 *   *
 6.    138 VVYMLHLATADVLFVSVLPFKISYYFSG           165  (SEQ ID NO:142)
19.     73 NYILLNLAVADLFMVLGGFTSTLYTSLH           100  (SEQ ID NO:143)
21.     61 NYFLLSLACADLIIGTFSMNLYTTYLLM            88  (SEQ ID NO:144)
30.     75 DVYLLNLALADLLFALTLPIWAASKVNG           102  (SEQ ID NO:145)
40.    145 NFFIVSLAVADLTVALLVLPFNVAYSIL           172  (SEQ ID NO:146)
                       |
           22222222222222222222222222222
           44444444445555555555666666666
           01234567890123456789012345678

TM3

                                    *
 6.    176 RFVTAAFYCNMYASILLMTVISIDR             200  (SEQ ID NO:147)
19.    111 NLEGFFATLGGEIALWSLVVLAIER             135  (SEQ ID NO:148)
21.     99 DLWLALDYVASNASVMNLLLISFDR              123  (SEQ ID NO:149)
30.    111 KVVSLLKEVNFYSGILLLACISVDR             135  (SEQ ID NO:150)
40.    183 KLWLTCDVLCCTSSILNLCAIALDR             207  (SEQ ID NO:151)
                                    |
           33333333333333333333333333
           22223333333333444444444445
           67890123456789012345678901
```

TM4

```
                                  *         *
 6.    215 TLGRASFTCLAIWALAIAGVVPLVLKE          241  (SEQ ID NO:152)
19.    149 GENHAIMGVAFTWVMALACAAPPLAGW          175  (SEQ ID NO:153)
21.    138 TPRRAALMIGLAWLVSFVLWAPAILFW          164  (SEQ ID NO:154)
30.    149 KRHLVKFVCLGCWGLSMNLSLPFFLFR          175  (SEQ ID NO:155)
40.    222 TVGRVLLLISGVWLLSLLISSPPLIGW          248  (SEQ ID NO:156)
                                  |
                    4444444444444444444444444444
                    3344444444455555555566666666
                    8901234567890123456789012345
```

TM5

```
                                  *  *       *
 6.    268 AYYFSAFSAVFFFVPLIISTVCYVSIIRC        296  (SEQ ID NO:157)
19.    201 ESFVIYMFVVHFTIPMIIIFFCYGQLVFT        229  (SEQ ID NO:158)
21.    186 PIITFGTAMAAFYLPVTVMCTLYWRIYRE        214  (SEQ ID NO:159)
30.    200 MVLRILPHTFGFIVPLFVMLFCYGFTLRT        228  (SEQ ID NO:160)
40.    267 RGYVIYSSLGSFFIPLAIMTIVYIEIFVA        295  (SEQ ID NO:161)
                                  |
                    5555555555555555555555555555
                    3333444444444555555555566666
                    6789012345678901234567890123
```

TM6

```
                                  *  *  *
 6.    313 FLSAAVFCIFIICFGPTNVLLIAHYSFL          340  (SEQ ID NO:162)
19.    252 RMVIIMVIAFLICWVPYASVAFYIFTHQ          279  (SEQ ID NO:163)
21.    365 RTLSAILLAFILTWTPYNIMVLVSTFCK          397  (SEQ ID NO:164)
30.    242 RVIFAVVLIFLLCWLPYNLVLLADTLMR          269  (SEQ ID NO:165)
40.    529 RTLGIIMGVFVICWLPFFLMYVILPFCQ          556  (SEQ ID NO:166)
                                  |
                    6666666666666666666666666666
                    3333344444444445555555556666
                    5678901234567890123456789012
```

```
TM7
                                    **  *
 6.    347 EAAYFAYLLCVCVSSISSCIDPLIYYYASSECQ      379  (SEQ ID NO:167)
19.    282 NFGPIFMTIPAFFAKSAAIYNPVIYIMMNKQFR      314  (SEQ ID NO:168)
21.    394 CVPETLWELGYWLCYVNSTVNPMCYALCNKAFR      426  (SEQ ID NO:169)
30.    281 NNIGRALDATEILGFLHSCLNPIIYAFIGQNFR      313  (SEQ ID NO:170)
40.    559 CPTNKFKNFITWLGYINSGLNPVIYTIFNLDYR      591  (SEQ ID NO:171)
                                    |
           77777777777777777777777777777777
           23333333333444444444555555555566
           90123456789012345678901234567890 1
```

A search using the amino acid sequence of SEQ ID NO:71 was performed against the HMM database resulting in the identification of a 7 TM receptor profile in the amino acid sequence of SEQ ID NO:71. The results of the search are set forth below.

```
Score:44.14     SEQ ID NO:71 aa24-191     HMM: aa1-174
                                                (SEQ ID NO:172)

     GNiLVIWvIcRyRRMRTPMNYFIvNLAvADLLFslft.MPFWMvYyvMq
  24 ANAWGILSVGAKQKKWKPLEFLLCTLAATHMLN-VAVPIATYSVVQLRR

     gRWpFGdfMCrIWmYFDYMNMYASIFfLTcISIDRYLWAICHPMrYmR
  72 QRPDFEWNEGLCKVFVSTFYTLTLATCFSVTSLSYHRMWMVCWPVNYRL

     WMTpRHRAWvMIiiIWvMSFlISMPPFLMFrWstyrDEneWNmTWCmIyD
 121 SNAKK-QAVHTVMGIWMVSFILSALPA-VG-W-HDTSERFYTHG-CRFIV

     WPewMWrWYvILmtiimgFYIPMiIMlF
 166 AEIGLGFGVCFLLLV-GGSVA-MGVICT
```

A search was performed against the HMM database resulting in the identification of a spectrin alpha chain profile in the amino acid sequence of SEQ ID NO:71. The results of the search are set forth below.

```
Score:8.78      SEQ ID NO:71 aa266-372    HMM: aa1-106
                                          (SEQ ID NO:174)

        IqeRMnElndRWerLkelMeqRRQMLedSmrlQQFfRDmDEeEsWInEK
    266 FSSLRADASAPWMALCVLWCSVAQALLLPVFLWACDRYRADLKAVREKC

        EqilnSDDYGkDLtsVQnLlkKHQaFEaDIaAHE.dRIqalnefaqqLIq
    315 MALMANDEESDDETSLEGGISPDLVLERSLDYGYGGDFVALDRMAKYEIS

        enHYasEe
    365 ALEGGLPQ
```

Tissue Distribution of STMST-1 mRNA

This Example describes the tissue distribution of STMST mRNA, as determined by Northern blot hybridization.

Northern blot hybridizations with the various RNA samples were performed (Clontech Multi-tissue Northern I and human fetal tissue northern) under standard conditions and washed under stringent conditions. A 4.5 Kb mRNA transcript was detected in heart, brain, placenta, lung, liver, skeletal muscle, fetal brain, fetal lung, and fetal kidney. Expression was highest in fetal brain.

Northern blot hybridization of poly A+ mRNA samples were also performed (Human Clontech poly A+ northern). A ~4.5 mRNA transcript was expressed in the following tissues at relative levels of heart > brain > placenta > liver > kidney.

EXAMPLE 14: CHARACTERIZATION OF STMST EXPRESSION BY RT-PCR

In this example, STMST expression levels were measured in a variety of tissue and cell samples using the Taqman™ procedure. The Taqman™ procedure is a quantitative, real-time PCR-based approach to detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA is generated from the samples of interest and serves as the starting materials for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) is included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-

4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. Table I sets forth the relative expression of STMST mRNA in a variety of tissues.

TABLE II: STMST Expression in Human Tissues

Tissue Source	Relative Expression	Tissue Source	Relative Expression
aorta/normal	0.39	liver/normal	0.99
fetal heart/normal	14.94	liver/fibrosis	1.14
heart/normal	5.70	fetal liver/normal	0.38
heart/CHF	3.23	lung/normal	1.30
vein/normal	0.22	lung/COPD	1.20
SMC/aortic	4.38	spleen/normal	0.64
nerve/normal	1.83	tonsil/normal	0.03
spinal cord/ normal	0.28	lymph node/ normal	0.18
brain cortex/ normal	25.30	thymus/normal	1.52
brain hypothalamus/ normal	20.47	epithelial cells/prostate	50
glioblastoma	0.75	endothelial	1.06

breast/normal	1.41
breast/tumor	1.40
ovary/normal	0.89
ovary/tumor	1.25
pancreas/normal	1.67
prostate/normal	1.87
prostate/tumor	2.12
colon/normal	0.31
colon/tumor	2.71
colon/BD	0.49
kidney/normal	1.12

cells/aortic	
skeletal muscle/normal	0.33
fibroblasts/ dermal	35.28
skin/normal	5.43
adipose/normal	0.10
osteoblasts/ primary	13.37
osteoblast/ undif.	0.75
osteoblasts/ diff.	0.75
osteoclasts	0.29

The highest expression was observed in epithelial cells, brain (cortex/hypothalamus), osteoblasts and dermal fibroblasts in the human tissue panel tested.

To further investigate the high expression in primary osteoblasts, STMST expression levels were measured by quantitative PCR using the Taqman™ procedure as described above. The relative levels of STMST expression in various cell lines is depicted in Figure 30A. The data demonstrate that at least three-fold STMST expression is seen in the ZB Osteo D18 cell line as well as in HUBCOB6 cells. Expression was also significantly increased in Clonetics D7 cells (*i.e.*, differentiated osteoblasts). The data presented in Figure 30B depict relative STMST expression levels in primary osteoblasts treated for 0, 1, 6 or 24 hours with either parathyroid hormone (PTH), interleukin-1 (IL-1) or dexamethasone (DEX). As clearly demonstrated by the data in Figure 30B, expression of STMST is upregulated in primary human osteoblasts stimulated for 24 hours with PTH. Transcriptional profiling analysis of a cDNA array (Figure 30C) confirms that expression in primary human osteoblasts is inducible by treatment of cells with parathyroid hormone (PTH). These data suggest that STMST and/or STMST agonism may mimic PTH anabolic effects on bone.

EXAMPLE 15: CHARACTERIZATION OF STMST EXPRESSION IN OSTEOGENIC CELLS BY NORTHERN BLOT ANALYSIS AND *IN SITU* ANALYSIS

Northern blot hybridization of poly A+ from the following samples was performed under standard hybridization and wash conditions: human bone (total mRNA), human bone (poly A+ RNA), HuBCOB6 (primary osteoblasts), HuBCOB11 (primary osteoblasts), huBCOB12 (primary osteoblasts), U2OS (osteoblast cell line), human spleen control (total mRNA) and human skeletal muscle (total mRNA). STMST transcript was also detected in human spleen mRNA.

In situ analysis was performed according to standard methodologies on tissue sections of human fetal bone. To generate a sense probe the following primers were used; forward primer: AGATGCCACCTTCCAGGCT (SEQ ID NO:85) and reverse primer: GGAGAAGTGCATGGCCCTC (SEQ ID NO:86) resulting in a sense probe having the following sequence: TCTCATCGTCTGACTCCTCGTCGTTGG (SEQ ID NO:87). Sense STMST probe hybridized particularly to osteoblasts within human fetal bone sections, consistent with coexpression of STMST with PTH-R positive osteoblasts. Importantly, this pattern of expression was very similar to the expression pattern for that of PTH-R, a good pre-osteoblast/osteoblast lineage marker.

EXAMPLE 16: STMST-2 mRNA *IN SITU* HYBRIDIZATION

This Example describes the characterization of brain-specific STMST expression.

The distribution of STMST-2 mRNA in mouse brain was examined as follows. Mouse brain was frozen with powdered dry ice, and cryostat sections were cut at 10 μm thickness through hypothalamus region, mounted on superfrost plus sides (VWR) and stored at -80° until needed.

Prior to analysis, mouse brain sections were air dried for 20 minutes and then incubated with ice cold 4% PFA (paraformaldehyde)/1xPBS for 10 minutes. The slides were then washed with 1 x PBS twice (5 minutes each time), incubated with 0.25% acetic anhydride/1 M triethanolamine for 10 minutes, washed with PBS for 5 minutes and dehydrated with 70%, 80%, 95% and 100% ethanol (1 minute each). Sections were incubated with chloroform for 5 minutes, rehydrated with 100% and 95% ethanol, then air dried. Hybridizations were performed using the following ^{35}S-radiolabeled ($5x10^7$ cpm/ml) cRNA probes: 5'GGCGGTGCACACAGTTAT'3 (SEQ ID NO:88) and 5'AGAGAGCGCTCCAAATACCAT3' (SEQ ID NO:89) in the presence of 50% formamide, 10% dextran sulfate, 1 x Denhardt's solution, 600 mm NaCl, 10 mm DTT, 0.25% SDS and 100 μg/ml RNase A in TNE at 37°C for 30 minutes, washed in TNE for 10 minutes, incubated once in 2x SSC at 60° for 1 hour, once in 0.2x SSC at 60° for 1

hour, 0.2x SSC at 65° for 1 hour and dehydrated with 50%, 70%, 80%, 95% and 100% ethanol. Localization of mRNA transcripts was detected by dipping slides in Kodak NBT-2 photoemulsion and exposing for 14 days at 4°C, followed by development with Kodak Dektol developer. Slides were counterstained with haemotoxylin and eosin and photographed. Controls for the *in situ* hybridization experiments included the use of a sense probe which showed no signal above background levels.

This analysis revealed that STMST-2 mRNA is expressed within the arcuate nucleus and the ventromedial nucleus of the hypothalamus, both of which are implicated in control of feeding behavior, as described herein.

EXAMPLE 17: REGULATION OF STMST EXPRESSION DURUING TUBE FORMATION OF ENDOTHELIAL CELLS

This example describes the characterization of STMST expression during tube formation of endothelial cells. Figure 31 shows relative mRNA expression levels of STMST, as determined by microarray hybridization. Human umbilical vein endothelial cells (HUVECs) were grown under a variety of conditions. On tissue culture plastic, expression of STMST is higher in endothelial cells than in non-endothelial 293 cells. When endothelial cells are cultured on Matrigel and form vascular tube-like structures, STMST expression is markedly down regulated. This result suggests that expression of STMST may inhibit vascular tube formation, so it is down-regulated to allow vascular tube formation to proceed. This process is believed to be similar to angiogenesis. In addition to its relevance to angiogenesis, the data also suggest roles in atherosclerosis and the control of vascular tone, as endothelial cell phenotype plays an important role in both of these processes. At least two genes with established relevance to atherosclerosis and control of vascular tone, cyclo-oxygenase-2 and endothelin-1, are regulated in this model.

EXAMPLE 18: EXPRESSION OF RECOMBINANT ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, and STMST PROTEIN IN BACTERIAL CELLS

In this example, ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST is fused to GST and this fusion polypeptide is expressed in *E. coli*, e.g., strain PEB199. As the human ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST protein is predicted to be approximately 18 kDa, and GST is predicted to be 26 kDa, the fusion polypeptide is predicted to be approximately 44 kDa, in molecular weight. Expression of the GST- ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST fusion protein in PEB199 is induced with

IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

EXAMPLE 19: EXPRESSION OF RECOMBINANT ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, AND STMST PROTEIN IN COS CELLS

To express the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST protein and an HA tag (Wilson et al. (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5a, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST -pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A*

Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST polypeptide is detected by radiolabeling (^{35}S-methionine or ^{35}S-cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labeled for 8 hours with ^{35}S-methionine (or ^{35}S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST polypeptide is detected by radiolabeling and immunoprecipitation using an ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, or STMST specific monoclonal antibody.

EXAMPLE 20: RETROVIRAL DELIVERY OF ITALY, Lor-2, STRIFE, TRASH, BDSF, LRSG, AND STMST INTO MICE

The entire open reading frame of ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, or STMST is subcloned into the retroviral vector MSCVneo, described in Hawley et al.(1994) *Gene Therapy* 1:136-138. Cells (293Ebna, Invitrogen) are then transiently transfected with the ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, or STMST construct and with constructs containing viral regulatory elements, to produce high titre retrovirus containing the ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, or STMST gene. This virus is then used to transfect mice. These mice are then tested for any gross pathology and for changes in their immune response using standard assays.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.